Molecular Cloning and Characterization of the 120-Kilodalton Protein Gene of *Ehrlichia canis* and Application of the Recombinant 120-Kilodalton Protein for Serodiagnosis of Canine Ehrlichiosis

**XUE-JIE YU, JERE W. McBRIDE, C. MARCELA DIAZ, AND DAVID H. WALKER**

Department of Pathology, WHO Collaborating Center for Tropical Diseases, University of Texas Medical Branch, Galveston, Texas 77555-0609

Received 11 December 1998/Returned for modification 2 April 1999/Accepted 24 August 1999

The 120-kDa outer membrane protein (p120) is a potential adhesin of *Ehrlichia chaffeensis* and recombinant p120 is very useful for serodiagnosis of human monocytotropic ehrlichiosis. The analogous gene of p120 in *Ehrlichia canis* was cloned, sequenced, and expressed. Like the *E. chaffeensis* p120, the *E. canis* p120 contains tandem repeat units. However, neither the repeat number nor the amino acid sequences in the repeats are identical in the two *Ehrlichia* species. The repeat units are hydrophilic and by probability analysis are predicted to be surface exposed in both species. The repeat regions of the p120s of the two species have common amino acid sequences that are predicted to be surface exposed. The overall amino acid sequence of the *E. canis* p120 is 30% homologous to that of *E. chaffeensis* p120. Protein immunoblotting demonstrated that the recombinant *E. canis* p120 reacted with convalescent sera from dogs with canine ehrlichiosis. These results indicate that the recombinant p120 is a potential antigen for the serodiagnosis of canine ehrlichiosis.

**Ehrlichia** spp. are obligate intracellular gram-negative bacteria which reside in the endosomes of hematopoietic cells and infect various animal hosts including humans, domestic and wild Canidae, deer, horses, sheep, cattle, and wild rodents. Each member of the tribe *Ehrlichiae* has its own particular target cell tropism. Most species of *Ehrlichia* are either monocytotropic (*E. canis*, *E. chaffeensis*, *E. sennetsu*, *E. risticii*, and *E. muris*) or granulocytotropic (*human granulocytic ehrlichia* [HGE], *E. equi*, *E. phagocytophila*, and *E. ewingii*) with the exceptions of *Cowdria ruminantium*, which grows in the endothelial cells of the host, and *Anaplasma marginae*, an erythrocyte parasite. Although ehrlichiae were described in the early part of this century, they were primarily considered pathogens of veterinary importance in the United States during this decade. Two new human *Ehrlichia* pathogens (*E. chaffeensis* and a human *E. phagocytophila*-like organism) were discovered in the United States (4, 6, 10, 17) recently. *E. canis*, the prototype species of the genus, is the etiologic agent of canine ehrlichiosis.

Canine ehrlichiosis is a worldwide disease transmitted by the brown dog tick, *Rhipicephalus sanguineus* (12, 16). *E. canis* causes a mild transient acute febrile illness, which may progress to severe illness and a fatal syndrome (tropical canine pancytopenia) (5, 11, 23). Each year millions of dollars are spent treating companion and working dogs infected with *E. canis* worldwide. Recently we demonstrated the *E. canis* p120 analogue exists in three North Carolina isolates (Demon, DJ, and Jake) was kindly provided by R. E. Corstvet (Louisiana State University, Baton Rouge). Ehrlichiae were cultivated in DH82 cells, a canine macrophage-like cell line (9). DH82 cells were harvested with a cell scraper when 100% of the cells were infected with ehrlichiae. The cells were dissolved with a Braun-Sonic 2000 sonicator at 40 W for 30 s twice on ice. The cell lysate was loaded onto discontinuous gradients of 42 to 36 to 30% Renografin and then centrifuged at 80,000 × g for 60 min. Ehrlichiae in the heavy and light bands were collected (24) and washed by centrifugation with sucrose-phosphate-glutamate buffer (218 mM sucrose, 3.8 mM KH$_2$PO$_4$, 7.2 mM K$_2$HPO$_4$, 4.9 mM glutamate, pH 7.0).

**DNA preparation.** *E. canis* genomic DNA was prepared from Renografin density gradient-purified ehrlichiae by using an IsoQuick nucleic acid extraction kit according to the instructions of the manufacturer (ORCA Research Inc., Bothell, Wash.). The genomic DNA was used in Southern blotting and in PCR for detecting the *E. canis* p120 gene.

**MATERIALS AND METHODS**

**Ehrlichia.** *E. canis* Oklahoma was kindly provided by Jacqueline Dawson (Centers for Disease Control and Prevention, Atlanta, Ga.). *E. canis* Florida and three North Carolina isolates (Demon, DJ, and Jake) were kindly provided by Edward B. Breitschwerdt (College of Veterinary Medicine, North Carolina State University, Raleigh). *E. canis* Louisiana was kindly provided by R. E. Corstvet (Louisiana State University, Baton Rouge). Ehrlichiae were cultivated in DH82 cells, a canine macrophage-like cell line (9). DH82 cells were harvested with a cell scraper when 100% of the cells were infected with ehrlichiae. The cells were centrifuged at 17,400 g for 20 min. The pellets were disrupted with a Braun-Sonic 2000 sonicator at 40 W for 30 s twice on ice. The cell lysate was loaded onto discontinuous gradients of 42 to 36 to 30% Renografin and then centrifuged at 80,000 × g for 60 min. Ehrlichiae in the heavy and light bands were collected (24) and washed by centrifugation with sucrose-phosphate-glutamate buffer (218 mM sucrose, 3.8 mM KH$_2$PO$_4$, 7.2 mM K$_2$HPO$_4$, 4.9 mM glutamate, pH 7.0).

**DNA preparation.** *E. canis* genomic DNA was prepared from Renografin density gradient-purified ehrlichiae by using an IsoQuick nucleic acid extraction kit according to the instructions of the manufacturer (ORCA Research Inc., Bothell, Wash.). The genomic DNA was used in Southern blotting and in PCR for detecting the *E. canis* p120 gene.

**CAROLINA E. canis** is the prototype species of the genus, is the etiologic agent of canine ehrlichiosis.

Previously we cloned and sequenced the p120 gene of *E. canis* (25). Very recently we demonstrated the *E. canis* p120 to be an outer membrane protein that is preferentially expressed on the dense-core ultrastructural form of *E. chaffeensis* but not on the reticular cell (19a). The p120 appears to be an adhesin of *E. chaffeensis* because a noninvasive, nonadherent strain of *Escherichia coli* expressing the p120 acquired the ability to adhere to and enter cultured mammalian cells (Popov et al., submitted). The p120 is an immunodominant protein of *E. chaffeensis*, and it reacts with sera from most patients with monocytotropic ehrlichiosis (27). *E. canis* and *E. chaffeensis* are genetically and antigenically closely related species (2, 3, 7). The homologies between *E. canis* and *E. chaffeensis* are 98% for the 16S rRNA gene and 89% for the nad4 gene (26). Since the p120 appears to be important in the attachment and serodiagnosis of *E. chaffeensis*, we hypothesized that an *E. chaffeensis* p120 analogue exists in *E. canis* and possesses similar biological functions. In this study we cloned, sequenced, expressed, and characterized the p120 gene of *E. canis* and evaluated the recombinant p120 of *E. canis* for serodiagnosis of canine ehrlichiosis by Western blotting.

**E. chaffeensis** is the prototype species of the genus, is the etiologic agent of canine ehrlichiosis.

Previously we cloned and sequenced the p120 gene of *E. chaffeensis* (25). Very recently we demonstrated the *E. chaffeensis* p120 to be an outer membrane protein that is preferentially expressed on the dense-core ultrastructural form of *E. chaffeensis* but not on the reticular cell (19a). The p120 appears to be an adhesin of *E. chaffeensis* because a noninvasive, nonadherent strain of *Escherichia coli* expressing the p120 acquired the ability to adhere to and enter cultured mammalian cells (Popov et al., submitted). The p120 is an immunodominant protein of *E. chaffeensis*, and it reacts with sera from most patients with monocytotropic ehrlichiosis (27). *E. canis* and *E. chaffeensis* are genetically and antigenically closely related species (2, 3, 7). The homologies between *E. canis* and *E. chaffeensis* are 98% for the 16S rRNA gene and 89% for the nad4 gene (26). Since the p120 appears to be important in the attachment and serodiagnosis of *E. chaffeensis*, we hypothesized that an *E. chaffeensis* p120 analogue exists in *E. canis* and possesses similar biological functions. In this study we cloned, sequenced, expressed, and characterized the p120 gene of *E. canis* and evaluated the recombinant p120 of *E. canis* for serodiagnosis of canine ehrlichiosis by Western blotting.

**MATERIALS AND METHODS**

**Ehrlichia.** *E. canis* Oklahoma was kindly provided by Jacqueline Dawson (Centers for Disease Control and Prevention, Atlanta, Ga.). *E. canis* Florida and three North Carolina isolates (Demon, DJ, and Jake) were kindly provided by Edward B. Breitschwerdt (College of Veterinary Medicine, North Carolina State University, Raleigh). *E. canis* Louisiana was kindly provided by R. E. Corstvet (Louisiana State University, Baton Rouge). Ehrlichiae were cultivated in DH82 cells, a canine macrophage-like cell line (9). DH82 cells were harvested with a cell scraper when 100% of the cells were infected with ehrlichiae. The cells were centrifuged at 17,400 g for 20 min. The pellets were disrupted with a Braun-Sonic 2000 sonicator at 40 W for 30 s twice on ice. The cell lysate was loaded onto discontinuous gradients of 42 to 36 to 30% Renografin and then centrifuged at 80,000 × g for 60 min. Ehrlichiae in the heavy and light bands were collected (24) and washed by centrifugation with sucrose-phosphate-glutamate buffer (218 mM sucrose, 3.8 mM KH$_2$PO$_4$, 7.2 mM K$_2$HPO$_4$, 4.9 mM glutamate, pH 7.0).

**DNA preparation.** *E. canis* genomic DNA was prepared from Renografin density gradient-purified ehrlichiae by using an IsoQuick nucleic acid extraction kit according to the instructions of the manufacturer (ORCA Research Inc., Bothell, Wash.). The genomic DNA was used in Southern blotting and in PCR for detecting the *E. canis* p120 gene.
Southern blotting. The E. chaffeensis p120 gene was amplified by PCR with primer pair pXcFb9 (CAG CAA GAG CAA GAA GTAG) and pXr4 (ACA TAA CAT TCC ACT TTC AAA). The 1.2-kb PCR product lacked 138 nucleotides at the beginning and end, of the structural gene of the E. chaffeensis p120. DNA was labeled during PCR by incorporating digoxigenin-dUTP with the PCR DIG probe synthesis kit (Roche Molecular Biochemicals, Indianapolis, Ind.) and was used as a probe to detect the homologous gene in E. canis. By Southern blotting, DNA hybridization was performed at 42°C overnight with the Dig Easy hybridization buffer, and the digoxigenin-labeled DNA bound to the E. canis genomic DNA was detected with nitroblue tetrazolium and BCIP (5-bromo-4-chloro-3-indolylphosphate) by following the instructions of the manufacturer (Roche Molecular Biochemicals). To ensure an appropriately representative partial digestion, an aliquot was removed at each of the 5' end of the insert before Southern blotting. The digoxigenin-labeled DNA probe was hybridized with the genomic DNA of E. canis p120. The DNA and deduced amino acid sequences were analyzed with the Wisconsin GCG software package (Genetics Computer Group, Inc., Madison, Wis.) and DNASTAR software (DNASTAR, Inc., Madison, Wis.). Both DNA strands of the E. canis p120 gene were sequenced. The DNA probes were labeled with [32P]dCTP by using a random-primer labeling kit (Perkin-Elmer Applied Biosystems, Foster City, Calif.).

Protein purification. The recombinant GST fusion protein was cleaved from the GST fusion vector (Qiagen Inc., Santa Clarita, Calif.) and was cloned into pCR2.1 TA cloning vector (Invitrogen, Carlsbad, Calif.). The resultant recombinant plasmid was designated pCR120.

Cloning the E. canis p120 gene. Primers were designed based on the DNA sequence of the E. chaffeensis p120 gene (Fig. 2). The E. canis p120 gene was amplified by PCR with 30 cycles of 94°C for 30 s, 52°C for 1 min, and 72°C for 2 min. The PCR product was purified by using a QIAquick PCR purification kit (Qiagen Inc.) and was cloned into pCR2.1 TA cloning vector (Invitrogen, Carlsbad, Calif.). The resultant recombinant plasmid was designated pCR120.

DNA sequencing. DNA was sequenced with an ABI Prism 377 DNA sequencing kit (Perkin-Elmer Applied Biosystems, Foster City, Calif.). Both DNA strands of the E. canis p120 gene were sequenced. The DNA insert was sequenced by using Exonuclease III with the Erase-a-Base system (Perkin-Elmer Applied Biosystems, Foster City, Calif.). Both DNA strands of the E. canis p120 gene were sequenced.

Expression of the E. canis p120 gene. Directly cloning the E. canis p120 gene into the pGEX expression vector (Amersham Pharmacia Biotech, Piscataway, N.J.) was prevented by the absence of matched restriction endonuclease-digestion sites, including the plasmid vector sequence. To ensure an appropriately representative partial digestion, an aliquot was removed at each of the 5' end of the insert before Southern blotting. The digoxigenin-labeled DNA probe was hybridized with the genomic DNA of E. canis p120. The DNA and deduced amino acid sequences were analyzed with the Wisconsin GCG software package (Genetics Computer Group, Inc., Madison, Wis.) and DNASTAR software (DNASTAR, Inc., Madison, Wis.). Both DNA strands of the E. canis p120 gene were sequenced. The DNA probes were labeled with [32P]dCTP by using a random-primer labeling kit (Perkin-Elmer Applied Biosystems, Foster City, Calif.). Both DNA strands of the E. canis p120 gene were sequenced.

RESULTS

Cloning the E. canis p120 gene. Southern blotting demonstrated that the E. chaffeensis p120 gene probe failed to hybridize with restriction enzyme-digested E. canis genomic DNA under conditions in which the probe gave strong hybridization with E. chaffeensis genomic DNA (Fig. 1). The control probe from the E. canis p120 gene hybridized with E. canis DNA but not E. chaffeensis DNA. These results indicated that the E. canis p120 gene differed substantially from the homologous E. chaffeensis p120 gene.

We further attempted to amplify the homologous p120 gene in E. canis Oklahoma by PCR. Primers derived from the E. chaffeensis p120 gene and sequences flanking the gene had been used previously for sequencing the E. chaffeensis p120 gene (Fig. 2). Three forward primers were paired with three reverse primers to form nine pairs of primers. A 2.5-kb DNA fragment was amplified from E. canis genomic DNA by the primer pair pXcFb9 and pXr3, derived from the noncoding DNA sequences flanking the E. chaffeensis p120 gene (Fig. 2). No DNA was amplified by using primers derived from the coding region of the E. chaffeensis p120 gene. The 2.5-kb PCR product was cloned into pCR2.1 TA cloning vector to generate the recombinant plasmid pCA120.
DNA sequence analysis of the *E. canis* p120 gene. Preliminary sequencing data indicated that the 2.5-kb PCR product of *E. canis* contained tandem repeats with 108 nucleotides each. The presence of the repeats made the sequencing difficult to accomplish by primer walking. Restriction enzyme analysis of the DNA sequences demonstrated that each repeat has a unique *Spe*I endonuclease cleavage site. Therefore, the number of repeats was determined by *Spe*I partial digestion and Southern blotting. Southern blotting demonstrated that there were 14 repeats in the *E. canis* p120 gene (Fig. 3). The repeat region of the *E. canis* p120 gene was sequenced by unidirectional deletion of the DNA fragment in pCA120. DNA sequencing demonstrated that the DNA insert contained an open reading frame (ORF) of 2,064 nucleotides which encoded 688 amino acids (Fig. 4). This ORF was designated the *E. canis* p120 gene. There were no consensus DNA sequences of the *E. coli* promoter near the 5' end of the gene. The N terminus of the deduced amino acids did not share consensus sequence with *E. coli* signal peptides. DNA sequencing confirmed that there were 14 repeats in the *E. canis* p120 gene (Fig. 4). At the amino acid level, the homology of all repeats was greater than 94% (Fig. 5). Preceding the first repeat there is an incomplete repeat that has a seven-amino-acid deletion (Fig. 5) and that is 70% homologous to the other repeats.

Sequence homology of the p120s of *E. canis* and *E. chaffeensis*. Searching the SwissProt database by using the Fasta program revealed that the amino acid sequence of the *E. canis* p120 is most closely related to that of the *E. chaffeensis* p120. The amino acid identity of p120s of *E. canis* and *E. chaffeensis* is 30%. A comparison of the amino acid sequences of *E. chaffeensis* and *E. canis* showed that they are more conserved on the N terminus and in the repeat region of p120.
acid identity is 50% for the first 32 amino acids of the N termini of the 120-kDa proteins of E. canis and E. chaffeensis.

The amino acid sequences of the E. canis and E. chaffeensis p120s, especially the repeats, were similar in hydrophobicity, surface probability, and antigenicity. All repeat units in both proteins are predicted to be hydrophilic, surface exposed, and highly antigenic (Fig. 6). The surface-exposed regions of the repeats have common amino acids in both the chrylicial species (Fig. 7).

Homologous genes in other strains of E. canis. A 2.5-kb DNA fragment from each strain of E. canis examined, including strains Florida, Louisiana, and the three North Carolina canine isolates (Demon, DJ, and Jake), was amplified with primers pxcf2 and pxar3. The segments of the p120 genes of all E. canis strains were sequenced on both the 5' and 3' ends. DNA sequence analysis demonstrated that the DNA sequences both up- and downstream of the repeat region were identical among all strains of E. canis. We did not attempt to sequence the complete repeat region for all E. canis strains because of the difficulty of sequencing the DNA repeats. We

FIG. 5. Phylogenetic relationships of the repeat units of the E. canis p120. The scale represents the percent difference in amino acid sequence.

FIG. 6. Surface probabilities, antigenic indices, and T-cell motifs of the p120s of E. canis and E. chaffeensis.

FIG. 7. Comparison of surface-exposed amino acids in repeat units of the p120s of E. canis and E. chaffeensis. (A) Surface probabilities of amino acids. Boldface letters indicate the amino acids conserved between E. canis and E. chaffeensis. (B) Alignment of the amino acid sequences shown in panel A. Lines represent identical amino acids. Dots represent conserved replacements. Dashes indicate gaps that were introduced for optimal alignment of the amino acid sequences.
sequenced the last repeats of all strains and the first repeat of DJ strain. The sequences of the first repeats of DJ and Oklahoma strains were identical. The sequences of the last repeats were identical among all strains. Homology of the p120 genes from all E. canis strains was further demonstrated by their identical SpeI restriction physical maps (Fig. 8).

Protein immunoblotting. The E. canis p120 gene was expressed in E. coli. The recombinant protein encoded by a 1,620-bp DNA fragment including all the repeats of the p120 gene was expressed as a GST fusion protein. The estimated molecular size of the fusion protein on sodium dodecyl sulfate (SDS) gel was approximately 140 kDa, which is much larger than the predicted molecular mass of the entire E. canis p120, which is only 73.6 kDa based on the amino acid sequence deduced from the DNA sequence (Fig. 9). Mouse antibodies to the recombinant p120 reacted with a p120 of E. canis (Fig. 9). The recombinant E. canis p120 reacted with all nine canine convalescent sera but with none of the normal dog sera (Fig. 10).

DISCUSSION

The homology of the amino acid sequences of the p120s of E. canis and E. chaffeensis is 30%. The DNA sequence homology of the p120 genes between the two species is 58%. It is surprising that the noncoding sequences flanking the p120 genes are more conserved than the coding sequences of the p120 genes of E. canis and E. chaffeensis. A comparison of 340 nucleotides upstream of the p120 gene revealed that the noncoding regions adjacent to the p120 genes of the two species of Ehrlichia have 84% homology. From an evolutionary point of view, the coding sequence that is under selection pressure would be expected to be more conserved than the noncoding sequence in which mutation would not be expected to affect the survival of the organism. We believe that the E. canis p120 gene is the homologue of the E. chaffeensis p120 considering that they are located in similar positions in the respective genomes, that they are 30% homologous, and especially that they have common motifs in the repeat region. The repeats in both proteins are hydrophilic and are predicted to be surface exposed. Even the total numbers of surface-exposed regions in the repeats of the two proteins are very close in spite of the difference in the numbers of repeat units (the E. chaffeensis p120 gene has three or four repeats, depending on the strain) (8, 25). The repeat units of both proteins have a common motif consisting of identical amino acids that are hydrophilic and that form the core of the surface-exposed regions of these proteins. These results indicated that the E. canis p120 is an outer membrane protein. The repeat units of both proteins are rich in glutamic acid and serine. Glutamic acid and serine each comprise 19% of the amino acids of the E. canis repeat unit. Glutamic acid and serine comprise 22 and 15% of the amino acids of the E. chaffeensis repeat units, respectively. Like that of the E. chaffeensis p120, the predicted molecular mass of the E. canis p120 is much smaller than the molecular size estimated on the basis of the electrophoretic mobility of the protein as determined by SDS-polyacrylamide gel electrophoresis (PAGE). The same phenomenon has been reported for other proteins containing repeat domains, including those of A. marginale (1), Plasmodium spp. (14), and Staphylococcus aureus (13, 20) and the HGE 100- and 130-kDa proteins (21). The repeat units of the HGE 100- and 130-kDa proteins have sequences in common with those of the E. chaffeensis p120 (21). The aberrant migration of the p120s of E. canis and E. chaffeensis is caused by glycosylation of the proteins (17a). Since the p120 of E. chaffeensis was differentially expressed in different ultrastructural forms of E. chaffeensis, this protein may play a role in the pathogenesis of E. chaffeensis infection. Whether or not the E. canis p120 is preferentially expressed in
the dense-core cell of E. canis is under investigation. The p120 gene appears to be conserved among all strains of E. canis since the known sequences, including the nonrepeat regions as well as the last repeats, are identical among strains of E. canis and since all E. canis strains have same number of repeats. The high degree of homology of DNA sequences and the identical numbers of repeats of the p120 genes among the strains of E. canis indicated that E. canis strains are genetically less diverse than those of E. chaffeensis, in which the number of repeats of the p120 gene differs among strains. p120 is immunodominant in both E. canis and E. chaffeensis because the recombinant p120s of both species react strongly with either human patient sera (27) or canine sera. Protein immunoblotting demonstrated that rabbit antisera to the E. chaffeensis p120 does not cross-react with E. canis and that mouse anti-E. canis p120 serum does not react with E. chaffeensis (data not shown). Therefore, the p120s of E. canis and E. chaffeensis may be useful for serodiagnosis of canine and human ehrlichiosis, respectively, for which they are both sensitive and specific.

ACKNOWLEDGMENTS

We thank Josie Ramirez-Kim for her assistance in the preparation of this manuscript.

This research was supported by a grant from the Clayton Foundation.

REFERENCES


